Quarterly Progress Report

for

Design and Packaging of Fault Tolerant Optoelectronic Multiprocessor Computing System

Sponsored by

Defense Advanced Research Projects Agency

Monitored by ONR Under Grant No. ONR/DARPA N00014-91J-1988

Grantee

The Regents of the University of California

University of California, San Diego

La Jolla, CA 92093

Reporting Period:

January 15, 1992-April 14, 1992

Principal Investigator:

Sing H. Lee

(619)534-2413

Program Manager:

William Miceli

(703)696-4715

1, Introduction

During this quarter, we continued our research efforts on (i) fabricating and testing packaged CMTM module, (ii) developing a CAD tool for solving PE placement problems in OE-MCMs, and (iii) designing fault masking and reconfiguration of the twin butterfly interconnection network. The results are summarized below.

2. Progress Summary

2.1 Opto-Electronic Packaging

A packaged version of a reconfigurable optical interconnection system known as the Correlation Matrix Tensor Multiplier (CMTM) was designed and the components were fabricated during the last quarters. In this quarter, we have assembled a packaged CMTM module and tested its capabilities of performing interconnections.

In the CMTM system, the two dimensional input array is correlated with an interconnection control tensor pattern to generate the output array. The control tensor pattern is designed to generated the desired interconnection. A random phase code is used for both the input and the control pattern to suppress the correlation output at the undesired positions. Fig.1 shows the schematics of the optical table setup of the CMTM system [1]. Fig.2 shows a packaged version of the CMTM module, with a phase code kinoform, two reflective CGH Fourier transform lenses and a $LiNbO_3$ crystal integrated on one face of the glass substrate. The design parameters of the first packaged version of CMTM, which uses visible light source, are shown in Fig.3. A photograph of the experimental packaging system is shown in Fig.4.

The input array is 5 x 5. The control pattern is 25 x 25, where each pixel of size 250 x 250 μm consists of 5 x 5 sub pixels of the phase code. The phase code and control pattern are transmissive plates, fabricated by e-beam lithography and reactive ion beam etching. The CGH Fourier transform lenses with an aperture of 12 x 12 mm were fabricated on a 1.5 mm thick glass plate. These lenses were designed by Code V for minimizing the spot diameter at the focal point for off-axis rays. The alignment of the relative position between two CGHs were achieved by e-beam writer with 0.1 μm accuracy. The photorefractive crystal is a 0.05% Fe-doped $LiNbO_3$ with dimension of 20 x 20 x 2 mm and oriented at c-axis perpendicular to the surface.

[1] *J. E. Ford, S. H. Lee, Y. Fainman, "Application of Photorefractive crystals to optical interconnection", Proc. SPIE, Vol.1215, pp155-165 (1990).*

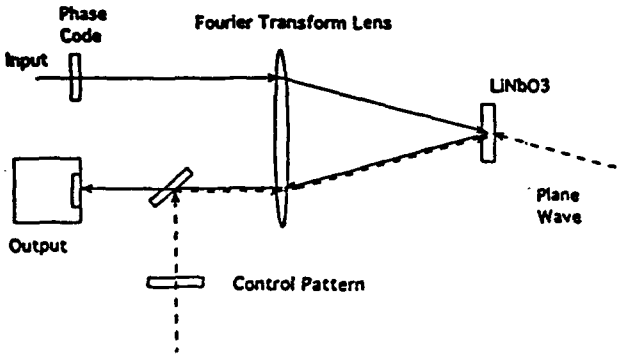


Fig.1. Schematic of the CMTM setup on an optical table.

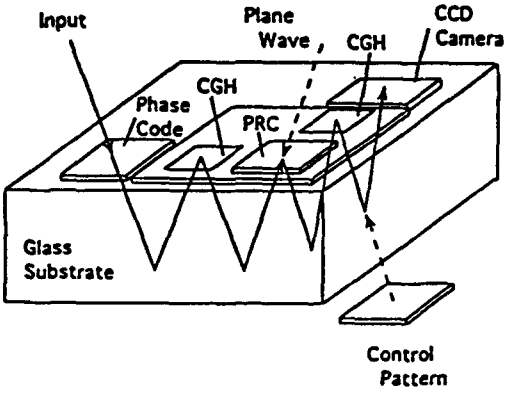


Fig.2. Schematic of the experimental packaged CMTM module.

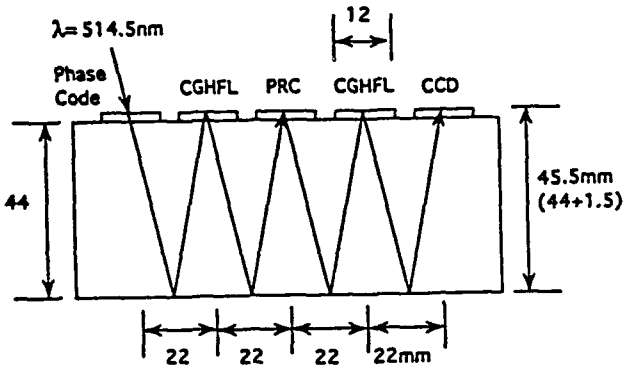


Fig.3. Design parameters of the CMTM packaging. CGHFL: CGH Fourier transform lens, PRC: photorefractive crystal.

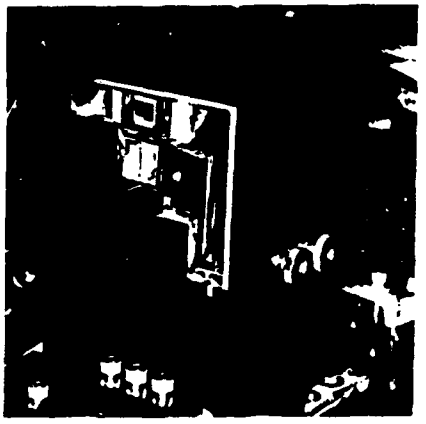


Fig.4. Photograph of the experimental packaged CMTM system.

The experiments were performed at 0.514 μm wavelength of an Argon laser. The maximum diffraction efficiency of the hologram recorded in the photorefractive crystal is 23% and the measured diffraction efficiency of the CGH of binary-level phase which was temporally used in the experiment is 32%. To determine the uniformity of the output signals at all the output locations, full interconnection between the 5 x 5 input and output arrays were examined to have SNR of 10.

From the CMTM packaging experiments we learned the following two important messages to improve the performance of the packaged system: (1) CMTM being space invariant requires CGH lenses of larger aperture (or areas) than the requirement from space variant systems, (2) Code V is indispensable in designing CGH lenses for their off-

axis operations. To keep the volume of packaged module small, our next packaged module will have lower f-# lenses; infrared sensitive photorefractive crystals will also be used so as to be compatible with diode lasers for compactness of the packaged module.

In addition to improve the CMTM system performance, research for the next quarters will concentrated on: (1) test of the application of Moiré method for alignment in packaging; (2) design and fabrication of holographic polarization selective beam splitters that are compatible with the packaging of reflective type SLMs; (3) design and fabrication of a photorefractive based space-variant optical interconnect module, whose schematics was shown in Fig.2 (b) of our former quarterly report (reporting period: Oct. 15, 1991 - Jan. 14, 1991).

2.2 Opto-Electronic CAD

New computer-aided design are needed for the development of OptoElectronic MultiChip Modules (OE-MCMs) utilizing free-space optical interconnects because the standard design used by electronics fail to properly model optoelectronic constraints. For instance, electronics minimize a cost function that incorporates the sum of all the interconnection distances, while optoelectronics minimize the maximum interconnection distance. We have been developing a new CAD tool which is useful for the placement of the processing elements (PEs) in OE-MCMs in order to minimize the maximum interconnection distance for large system size. In this quarterly report we report our results of computer simulations on placement of PEs in OE-MCMs based on a simulated annealing algorithm. Fig.5 shows the CGH interconnection model which was used as a basic configuration in the development of our placement algorithms. The maximum off-axis angle θ which determines the maximum lateral distance between two PEs is limited by the minimum feature size of the off-axis multi-level phase CGHs. As the number of PE increases and the fanout from (and to) each PE increases and becomes irregular, the maximum d becomes increasingly difficulty to minimize. The objective of our placement algorithm is to minimize the maximum value of d by optimizing the placement of PEs on each PE plane.

Fig.6 shows a schematic diagram of the physical model for a multi-stage OE-MCM. In this physical model, we assume there exist N (for an N-1 stages network) PE planes, that are interconnected to adjacent PE planes by multiple optical links. Each link in considered to be of the configuration shown in Fig.5 and thus the CGH planes are arrays of multiply superimposed off-axis diffractive lenses. We assume that each link begins at a source (modulator or laser diode) and terminates at a detector. Fig.7 shows the logical diagram of the placement problem. The interconnection pattern between a set

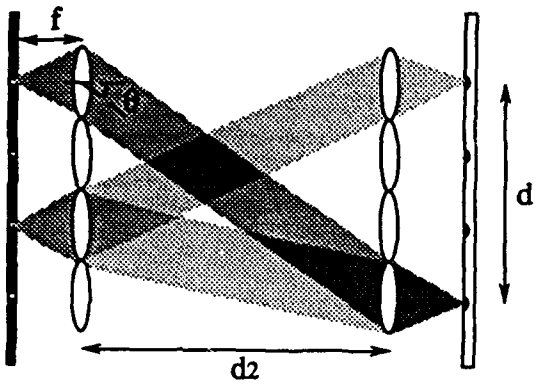


Fig.5. Schematic of CGH interconnect. Arbitrary interconnection design, one-to-one, one-to-many and many-to-one interconnects, can be obtained.

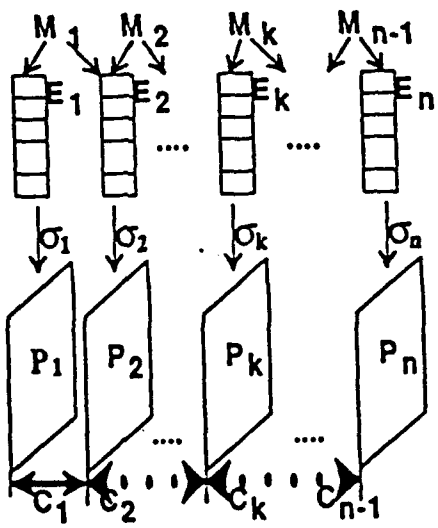
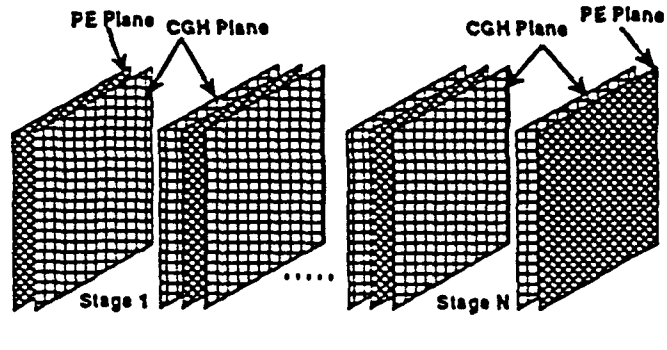
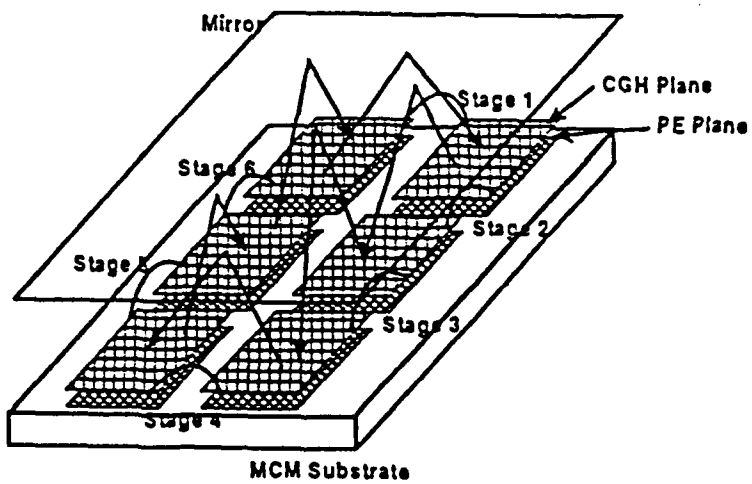


Fig.7. Logical relationship inside a general multistage interconnection network system.



(a)



(b)

Fig.6. OE-MCM physical model. (a) Transmissive configuration. (b) Reflective configuration.

of logical PEs (e_{kr}) in set E_k and those in sets E_{k-1} and E_{k+1} is determined by a netlist which provides the interconnect topology. The placement of these logical PEs into the actual physical array of P_k by the mapping σ_k determines both the interconnection distance and direction. We define the cost function C_k to be the maximum interconnection distance from each plane (say P_k) to any adjacent planes (P_{k-1} and P_{k+1}),

$$C_k = \max_{ij} \{d_{ki(k-1)j}, d_{ki(k+1)j}\}$$

The goal of our placement algorithm is then to minimize this cost function over all stages, $k = 1, 2, ..., N-1$. Simulated annealing algorithms has been used to solve this problem. The simulated annealing is a well-known algorithm used to solve problems that commonly occur in combinational optimization. In order to use simulated annealing, the problem is first configured as a system with a discrete number of states. The cost function C_k is used to describe some feature [i.e., $\exp(-\Delta C_k/T)$] about the system that the algorithm is to minimize. The state periodically evolves by reducing the temperature T which is used to calculate the value of $\exp(-\Delta C_k/T)$ for determining whether a positive change in C_k is acceptable or not; that prevents simulated annealing from getting caught in a local minimum.

The simulated annealing algorithm was applied to the placement problem for a real twin butterfly design example. The results are shown in Fig.8. For comparison, the result of using straight-forward placement, i.e. raster order placement of all PEs into the PE planes is also presented. The longest interconnection distance in straight-forward placement is 8.60, while that used simulated annealing algorithm is 4.24. In other words, the longest interconnection distance can be reduced by 50%, and therefore the distance between CGH planes as well as the volume of the twin butterfly network can be reduced by 50%. If the complexity of CGH fabrication were to be maintained at maximum for the same interconnection distance and network volume, the size of the network can be increased by a factor of 2.

Next Quarter, we plan to design and fabricate CGH arrays for the twin butterfly network based on the results obtained by the placement algorithm. In addition, we will

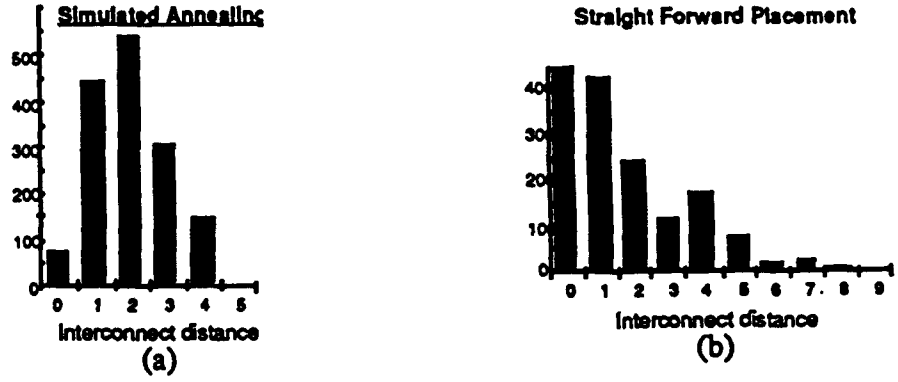


Fig.8. Histogram of the interconnect distances for the placement of a 64 node, 6-stage twin butterfly using (a) simulated annealing algorithm placement., (b) straight-forward placement

investigate another algorithm that comes from computer science, i.e. Matching algorithm to see if the computing time can be reduced. The result will be reported next quarterly report.

2.3 Fault Tolerance and Testing in Opto-electronic Computing

In the previous report we described the stuck fault model and design modifications to support parallel optical testing of the fabricated optoelectronic chip. This report will focus on the fault masking and reconfiguration of the twin butterfly interconnection network.

The essence of testing is controllability and observability. While it is straight forward in concept to drive a probe pad to the desired state during testing and observe it through another probe pad, how to do that in a packaged optoelectronic system is much more challenging. To achieve parallel testing, we modified each switching element (SE) (see Fig.9), to provide the test patterns that would be used to test its neighbors. Simple XOR gates can be used as 1-bit comparators to verify detector input against the test pattern. When there is a disagreement, a flag would be set to mask this detector out of active service. This is the approach used in the forward testing and backward testing. We have equated a modulator fault with four detector faults, i.e., a modulator fault is perceived by its neighbors as four detector faults (on four separate switching elements). In the event that all four input detectors of a switching element has been diagnosed as faulty, the entire switching element will be pronounced dead so that it can be removed from active use.

The forward testing procedure is summarized in the following algorithm:

step 1: *assign test pattern to be "0"*
step 2: *output test pattern through datamod*
step 3: *compare each datadet with test pattern,*
if disagree then mark that datadet as faulty
step 4: *if all datadet's are faulty, mark SE as faulty*
step 5 through 8: repeat step 1-4 with test pattern="1"

In backward testing, we need at least one detector to communicate reliably with each half of the next stage, i.e., at least two of the four handshaking detector must be functioning to keep this switching element useful. This is reflected in the following algorithm:

step 1: *assign test pattern to be "0"*
step 2: *output test pattern through CTSmod*
step 3: *compare each CTSdet with test pattern,*

if disagree then mark that CTSdet as faulty

step 4: *if both CTSdet leading to the upper half destinations are faulty or both CTSdet leading to the lower half destinations are faulty, then this SE cannot route message reliably, mark SE as faulty*

step 5 through 8: *repeat step 1-4 with test pattern="1"*

The results of both forward and backward testing are stored in registers (data det dead and CTS det dead, see Fig.10). A result of "1" indicates presence of fault and removes the corresponding detector from future use. While these testing procedure allow individual switching elements to detect and mask out faults, it is not apparent how a faulty switching element can be avoided during routing operation. This can be explained by the help of Fig. 10. Our approach is to propagate the status of the switching element during backward testing. The combinational logic right before the CTS modulator (low left corner of Fig.10) will force the modulator output to logic 0 if the corresponding data detector is faulty or the entire switching element has been declared useless. After one iteration of backward testing, the status of the $(log_2N)th$ stage switching elements would have propagated to the $(log_2N-1)th$ stage. Network reconfiguration is thus achieved by ignoring links leading to faulty switching elements. We will simply run the system test

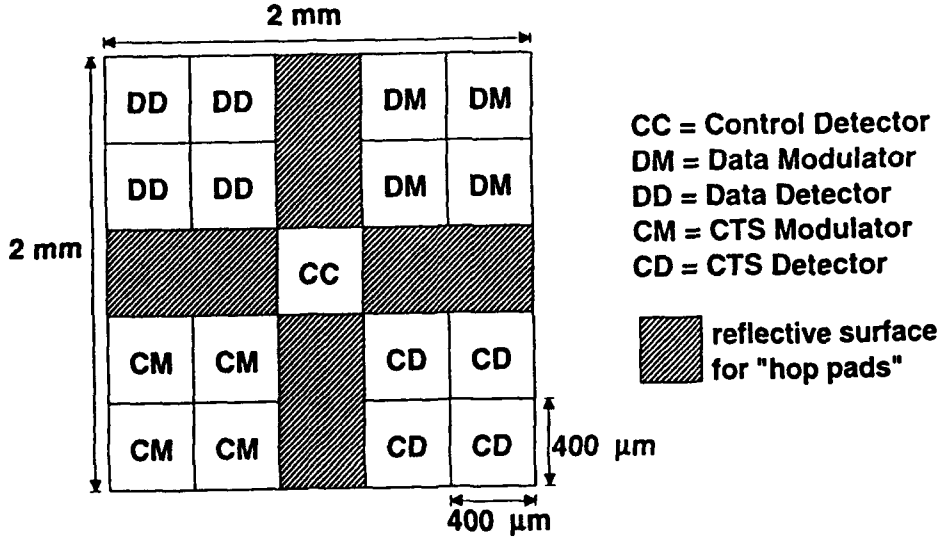


Fig. 9. Layout of interconnection CGH for each switching element. This CGH is packaged above the layer consisting of silicon logic and optoelectronic devices. The modulator outputs are fanned out to four "random" destinations by the larger CM and DM facets. The remaining facets focus incoming beams onto the detectors on the devcice layer. We could potentially use the unused area as "hop pads" for long interconnections that cannot be reached due to angular deflection constraints imposed by the minimum feature size used.

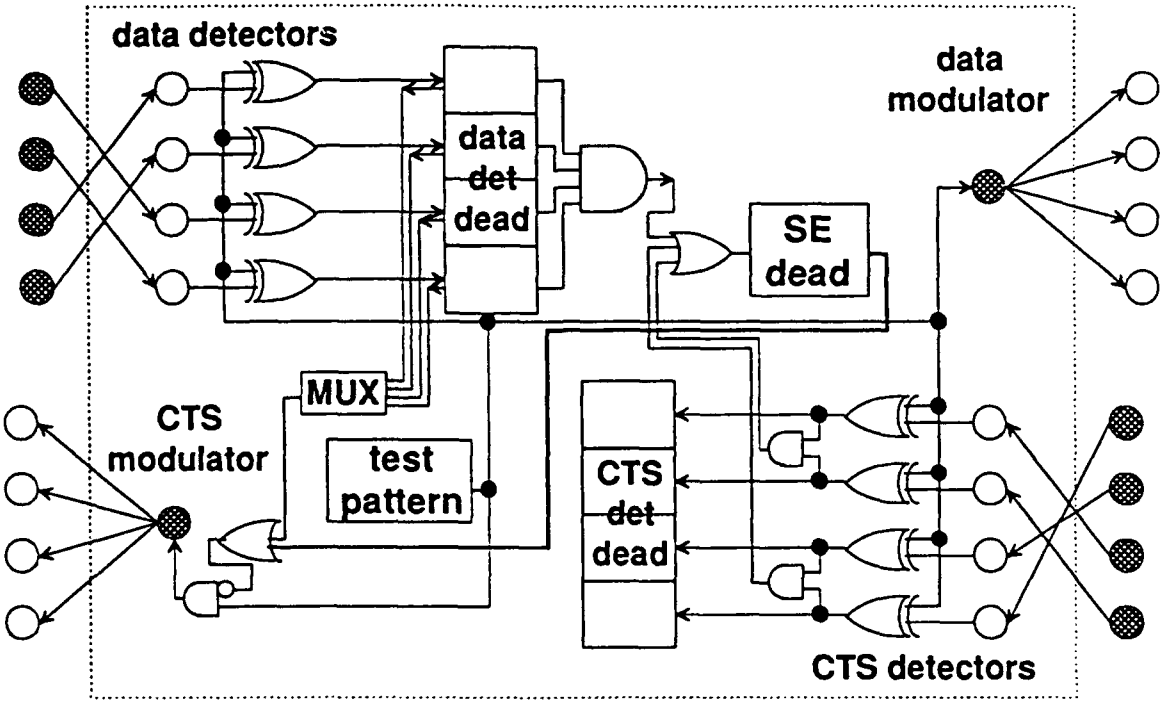


Fig. 10 Switching logic additions to support system testing and reconfiguration. Dotted lines mark the boundaries of a switching element. During testing, each switching element compares the test pattern it is sending (from *test pattern* via modulators) against the pattern it is receiving. The XOR gates (i.e., 1-bit comparators) indicate if there is disagreement and remove the faulty device from service by setting the corresponding flag in *data det dead* or *CTS det dead*. Once device failures are serious enough to render the whole swithcing element unreliable, it sets *SEdead* register to "1". This will make CTS modulator appear as *stuck-at-0* to the preceeding switching element, effectively removing itself from service. It takes $\log_2 N$ steps to propagate the status of the switching elements from the last stage to the first stage, therefore $\log_2 N$ steps is required to complete the reconfiguration.

for $log_2 N$ iterations to reconfigure the entire $log_2 N$ stage network without having to alter the interconnection CGH. If there are so many faults in the network that some input are prevented from reaching the output, another interconnection hologram would be needed.

We have analyzed the reliability of this switching element design with a combinatiorial reliability model, assuming exponential failure law for the optoelectronic devices. This law predicts a failure rate that would remain constant throughout the useful life of the component after an initial burn-in period. The result of the reliability analysis has been summarized in Table.1, assuming that we have empirical data for detector and

modulator failure rates. Table.1 could be used in a top down fashion, i.e., given an application with a certain reliability/availability requirement and derive the quality of the devices that would be needed to meet the specification.

Besides continuing researches on the parallel testing in the next quarter, we are going to compare VLSI and optoelectronic twin butterfly with large grain size switching elements, in order to determine when an optoelectronic implementation would be preferred over a pure VLSI implementation.

Table 1. SE Reliability Calculations

			Remarks
SLM failure rate	λm	0.000200000	failures/hour
detector failure rate	λd	0.000010000	failures/hour
time	t	100.00	hours
SLM reliability	Rm	0.980198673	$\exp(-\lambda mt)$
detector reliability	Rd	0.999000500	$\exp(-\lambda dt)$
R(data detectors)	Rdd	1.000000000	$1-(1-Rd)^4$
R(CTS detectors)	Rcd	0.999996008	$1-(1-Rd^2)^2$
SE reliability	Rse	0.960785604	$Rm^2*Rdd*Rcd$
SE failure rate	λse	0.000400040	failures/hour
		0.400039920	failures/1000 hr
Mean time to failure	MTTF	2499.75	hours